

大数据背景下高校档案管理模式变革研究

□ 贵州师范大学档案馆 姜 萍 厦门大学经济学院 姜叶飞

摘 要 大数据时代,高校档案管理工作面临新的挑战,传统档案管理模式已不能满足时代的需求。利用大数据海量存储的特点,可以提升档案在存储、管理、利用等方面的功效,进而为学校决策、社会利用提供快捷方便的资讯服务。

关键词 大数据 档案管理 模式 功能

中图分类号 G271 **文献标识码** A **收稿日期** 2014-9-9

随着国家信息化建设的不断深入,由云计算、云存储等新型的IT服务模式所直接催生的数字信息正在爆炸式增长,各高校档案馆(室)数据资源总量也在迅速增长。面对数量巨大、异构的文档信息,同时还要与时俱进地满足现代人类对于信息多样化、高效化、个性化、专深化的需求^[1],传统的档案管理模式显然已经不能满足时代的需求。大数据时代在不断改造传统产业的同时,也对传统档案管理模式提出了新的挑战^[2],档案管理理念及管理模式变革在所难免。笔者认为,尤其是在知识经济时代,大数据时代的来临相比其他信息技术更加契合档案管理工作的需要,档案资源转化为知识资源,将成为今后高校档案管理工作必然趋势,大数据在档案管理中将起到举足轻重的作用。由此可见,在大数据时代下,高校档案管理模式如不及时进行转变,高校档案管理工作将会“失能”,未来的档案馆(室)也将面临边缘化的危险。

一、大数据导致高校档案管理从量变引起质变

大数据时代下的数据处理和存储不仅越来越快,越来越方便,还越来越便宜。海量数据由结构化向非结构化转变,非结构化数据包括所有格式的办公文档、文本、图片、XML、HTML、各类报表、图像和音频或视频信息等。档案系统下存储的数据更多的是文本、图像、影音等,属于非结构型数据,通过大数据技术,对档案信息进行充分和全面的整合,使之有效地服务于学校及社会。

未来的大数据将直接成为决策的基础。过去的决策靠的是管理者个人的经验、直觉和智慧,但个人的经验毕竟是有限的,科学合理的决策要依靠数据说话。档案作为原始记录,由于其全面性、真实性、现实性等特点,使之具有其他任何材料不可替代的参考和凭证作用。通过对海量档案数据进行分析得出的决策,可以完整、客观地反映真实需求,从而作出正确的决策。因而,在大数据时代,我们完全可以只需要知道要做什么,而不需知道为什么要这么做。不再热衷于寻找因果关系,而关注其存在的相关关系,直接告诉我们“是什么”而不是“为什么”,这就导致档案管理从量变到质变,管理方式也将因此产生变革。用数据说话、用数据管理运行、用数据进行创新的大数据决策成为不可逆转的历史潮流。高校档案馆(室)作为学校档案数据存储部门,管理职能将发生根本性变化,不仅存储数据,还需要对数据进行分析,并将数据分析结果汇编成各种专题及基础数字汇编,为学校领导决策提供依据,为社会提供咨询服务,充分发挥档案史料的“资政”作用。

二、大数据导致高校档案管理技术层面的更新

在数字信息技术环境下,电子载体和数字传播方式的异军突起、信息处理方式的变革以及电子文件的爆炸式增长使得档案的传统载体形态开始向数字形态过渡^[3]。因而,构建数字档案馆(室)是数字时代或者说大数据时代档案馆(室)的发展方向。在目前强大的计

算机硬件和数据库的支持下,高校档案馆(室)既要按照“存量档案数字化”的要求,对馆藏传统载体档案进行数字化,把纸质文档通过扫描、导入到计算机数据库,还要按照“增量档案电子化”的要求,将原生电子文件导入计算机数据库,以便计算机存储档案信息。“存量档案数字化”与“增量档案电子化”是未来高校数字档案馆数字资源的基础,具有检索快捷方便、容易实时同步备份、存储空间小、维护方便、安全等效果。

随着数字档案馆建设的不断推进,未来产生的档案数字资源将是非常庞大的数据量,如果仅依赖于目前的磁性和光学介质进行备份会有很大的局限性,由此推进档案云备份技术的运用,即将异质备份转变为云备份。在这个前提下,高校档案馆(室)不需要花太多资金和人力建设大机房,直接购买服务就行,同时也不用担心在公网上的信息泄露。

三、大数据时代推动高校档案管理模式的变革

1.对高校归档范围和保管期限提出挑战。大数据时代是一个存储介质廉价的时代,数据的价值远远高于介质的价值。在当今“大档案”的管理理念下,传统的归档范围显然已不适用,大数据却为我们提供庞大的存储空间。为满足未来的需求,任何数据都可纳入归档范围。作为数据存储地的高校档案馆(室),档案数据可以包罗万象,不受归档范围与保管期限的限制。就一个学生来说,高校档案馆(室)可以利用大数据海量性的特点,可将其在校期间学习、生活、娱乐等各方面的信息纳入归档范围,甚至包括学生的考试试卷、绘画作品、各种便条等,做到应存尽存。同时,为满足学校及社会对档案信息的需求,档案“保管期限”概念将会变得模糊,任何档案价值都有可能成为永久。档案管理者不用投入过多的时间去研究、界定某份材料的价值,不再区分短期、长期、永久,统统以永久期限存储。因为没有任何人能断定30年后某个学生会是否能成为有影响的管理者、专家或学者。将来某个学生当年的只言片语极有可能成为有价值的珍贵档案,他的某一幅作品也很有可能成为价值不菲的宝贝。

2.对高校档案实体分类提出挑战。大数据时代数据处理和存储不仅越来越快,而且越来越方便,这对传统档案的分类、编目、排架等提出挑战。可以说,在大数据时代,现行高等学校档案实体分类将被取消,而改变为以时间为序的储存方式,因为时间序列区别程度强于分类归档。加之现有检索技术水平,不需要按照过去的分类建立索引,更不需要分类储存。就高校档案实体而言,其基本职能活动所形成的党群、行政、教学、科学研究、基本建设、仪器设备、产品生产、出版物、外事、财会等十个门类档案,不必按一级类目、二级类目细分,存储上也没必要按照分类排列,只需以时间按照书架顺序排列即可。档案编目可以书架序号来编制档号,在电脑上直接索引书架序号或档号即可找到需要的档案。检索工具也将变得更加科学、

基于电子病历系统的高血压病案数据挖掘研究

□ 辽宁中医药大学信息工程学院 孙艳秋 沈阳师范大学教育技术学院 刘 钢

摘 要 随着医院信息化的发展,以电子病历系统为载体的、记录治疗各种病的档案数据也越来越多。本文分析了基于电子病历系统的数据挖掘的特点,并引入到高血压病人档案的研究中,为高血压病案研究开辟了新的方向。

关键词 电子病历系统 高血压病案 数据挖掘

中图分类号 G271 **文献标识码** A **收稿日期** 2014-12-3

高血压病案是病人在医疗活动中形成的,它包括高血压的中医症候、症状、中药、中成药以及西医的诊断、西药用药的临床记录与总结。而基于电子病历系统的高血压病案不是像大众所认为的用计算机书写病历,而是高血压病人整个信息的总集成,包含了原纸张病历的所有静态信息,以及其他由于数字化带来的各种功能扩展和相关服务(影像资料、多媒体动态录像、诊断知识库、药品知识库、查询功能等),在电子病历上表现是高血压病的文字、表格、数字、图片、声音、影片等多种数据格式的完整统一,是一个详细记录高血压病的多媒体临床医疗数据库。收集、整理和利用高血压病案,可以将中西医高血压病的诊疗经验得以顺利传承,应用现代科学技术对这些高血压病案进行科学解析,特别是进行信息化、数字化和知识化的研究,不仅能丰富中医药学的理论体系,还能对整个医学科学的发展产生巨大的推动作用。

数据挖掘技术是目前促进中医学信息化的有效途径之一,它可以较好地发现其中的医学诊断规则和模式。本文通过对高血压病案

的临床诊疗进行数据挖掘和分析,将分析高血压的病因、症候、用药规律,以及它所带来的并发症研究。尤其通过对高血压病案的数据挖掘,全面解析该病的规律,分析该病个体化诊疗信息特征,以及大众病症特点,提炼出治疗该病的新理论、新方法、新知识,实现病症总结的经验和传承。但是随着高血压病例的增多,随着纸质档案的增多,进行数据挖掘的难度也在增加,尤其是录入数据库当中的时候,数据不同的表示方式带来了不同的数据结果。而随着信息化的发展将会产生越来越多的数据,电子病历系统的应用,使得传统的病案变成了数字化的病案,所以在电子病历系统中分析和挖掘这些数据,不但可以获得新知识、了解新规律,还可以研究新的算法,可以给高血压病案的研究带来新的发展和机遇。

一、电子病历系统和数据挖掘

所谓电子病历系统就是指医院内部支持电子病历信息的采集、存储和访问,并围绕提高医疗质量、保障医疗安全、提高医疗效率而提供信息处理和智能化服务的计算机信息系统。既包括应用于门

规范,提供利用也变得更为快捷、方便。

四、大数据时代提升高校档案管理功效

1. 知识发现和数据挖掘辅助学校决策。在当今知识经济时代,高校档案馆(室)的职能将发生根本性的改变,不只是保管,更为重要的是将档案信息转化为知识,档案管理者不仅是管理员,更是知识的提供者。档案资源转化为知识资源,将成为今后高校档案管理工作发展的必然方向,大数据在此将起到举足轻重的作用。对于馆藏档案可运用数据挖掘技术,例如关联规则、决策树、神经网络、支持向量机、随机森林等方法,可以从过往无用的“垃圾数据”中发现其中存在的规律,为学校领导层提供决策依据。如学生档案分析,可以区分和预测学生的类型和行为,通过学生基本信息、考勤、成绩、宿舍和图书馆出入记录、一卡通消费等数据的分析,可以预测哪些学生自杀概率较高,以便及时疏导,通过学生考勤、一卡通及其关联账户的使用情况,区分哪些学生是属于生活困难型的,从而更好地分配勤工助学岗位。同样地,学生作弊概率、健康状况、科研工作能力等都可以通过档案数据分析得出。在学校层面上,通过对网络、BBS、报刊等外部数据以及学校自身、教职工等档案数据及时预测学校可能发生的重大问题,找出哪些教师有较高的教学质量和科研水平,区分受欢迎和不受欢迎的课程以便改革,通过教师专业结构档案的数据分析,有助于了解高校学科分布的情况,为学校引进人才提供参考,通过对学校科研数据的分析,可以辅助教师申报国家自然科学基金、社会科学基金等各类课题。

2. 档案由“国家资源”到“社会资源”转变。在大数据时代下,各行各业的档案信息变得更为开放,档案数据逐渐趋于一体化。存储

资源管理、数据分类、重复数据删除、持续数据保护、归档、档案跟踪、备份等终将达到档案数据资源的最佳管理和开放利用效果。近年来,各省、市都相继建立区域性的档案管理网络,使某一地区、某一行业或某一系统的档案实现资源共享。档案由“国家资源”到“社会资源”,这是档案事业的一项重大战略转向,是一场触及思想观念、管理模式的改革,也是档案信息整合与利用模式的改革^[9]。高校档案馆(室)的主要优势在于储存、收藏了大量的原始档案,这些都是国家的重要信息资源,是社会资源中最为精华的部分。为此,高校档案馆(室)要积极转变观念,对可公开档案及解密文件等开放公开接口查询,从而实现“国家资源”到“社会资源”转变。这种开放利用既可带动社会进步,使各行各业建立互通互补的档案信息资源,同时打破传统馆藏资源概念,为社会利用提供更为快捷方便的资讯服务。由此而来,高校档案馆(室)服务社会的功能将进一步提高,高校档案的价值与作用也将得到进一步凸显。

总之,大数据时代,通过大数据技术、档案云技术等实现高校传统档案管理模式的变革及管理职能转变,档案管理人员也将从繁琐的传统工作程序中解脱出来,腾出更多的精力和时间对馆藏资源进行分析、研究及开发利用,从而降低管理成本,提高管理效率,达到档案管理服务学校决策、服务社会的目的。

参考文献

- [1] 大数据时代来临,如何提升档案价值 [EB/OL]. [2012-07-20] <http://www.pconline.com.cn/pcedu/344/3446656.html>.
 - [2] 杨冬权.在全国数字档案馆(室)建设推进会上的讲话[J].中国档案,2013(11).
 - [3] 裴友泉.档案信息由“国家资源”到“社会资源”重大战略转向研究[J].档案学,2012(6).
- ★作者姜萍为贵州师范大学档案馆馆长,研究馆员,姜叶飞为厦门大学经济学院统计系2011级硕士研究生。